Analysis of Distance Transforms for Watershed Segmentation on Chronic Leukaemia Images

T. Ahmad Aris, A. S. Abdul Nasir and W. A. Mustafa
Faculty of Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus,
Sungai Chuchuh, 02100 Padang Besar, Perlis, Malaysia
sqeifa.aris@gmail.com

***Abstract*—Leukaemia is a blood cancer that contributes to the increase in the world mortality rates per year. Leukaemia can be divided into two major types which are acute and chronic leukaemia. This disease is caused by the excessive production of abnormal white blood cells (WBCs); hence these cells play a major role in the screening and diagnosis of leukaemia disease. Leukaemia screening requires the complete blood count process. However, due to the cells complex nature in chronic leukaemia which is overlapped, it would be difficult to obtain the accurate number of the WBCs for the screening process. Therefore, this paper proposes an automated WBCs counting with analysis of watershed segmentation for the screening of chronic leukaemia images. The segmentation approach consists of a few steps; (1) colour conversion, (2) image segmentation, (3) noise removal and (4) separation of overlapping WBCs. In this paper, three different distance transforms for watershed segmentation known as Euclidean, city block and chessboard have been analysed in order to find the best approach which is capable of separating the overlapping WBCs. The experimental results show that segmentation using watershed based on Euclidean has successfully segmented 50 blood images with average counting accuracy of 99.81%, as compared to the city block (91.09%) and chessboard (98.78%). Thus, the proposed procedures with watershed segmentation provide an efficient alternative in enhancing the accuracy of the WBCs count for leukaemia screening.**

***Index Terms*—Chronic Leukaemia; Distance Transform; Otsu's Thresholding; Watershed Segmentation.**

I. INTRODUCTION

Cancer is one of the most common diseases in the world [1], [2]. This disease is created by abnormal cell growth excessively. There are more than hundred different types of cancer, and one of them is leukaemia. Leukaemia is defined as cancer that starts from blood-forming cells in the bone marrow. This disease is the 6th most common cancer in general Malaysia population [3]. Leukaemia can be divided into two major types which are acute and chronic leukaemia. Chronic leukaemia can be described as slow-growing cancer. There are two main types of chronic leukaemia which are chronic lymphocytic leukaemia (CLL) and chronic myelogenous leukaemia (CML), where CLL is the most common leukaemia in the adult. In leukaemia, the cell will stop developing normally when it became leukaemia cell. Over time, leukaemia cell will dislocate the normal blood cells. This can lead to serious problems such as anaemia, bleeding and infections.

Leukaemia can be cured if it is treated at the early stage. The procedure for early detection of leukaemia is called screening test. Generally, in leukaemia screening process, haematologists will look for the large number of abnormal WBCs on the stained slide and performs the complete blood count process. In this process, manual counting under the microscope is one of the methods that is still used for blood cell counting [4]. However, the possibility of faulty detection due to the human error may arise [5]. As a result, various applications of image processing techniques to segment the WBCs [6-8] and performing the WBCs counting process [9-11] have been reported for leukaemia screening.

For instance, Harun *et al.* [10] proposed a computer-aided system to perform the automated WBCs counting in normal and acute leukaemia images of 10X magnification. This study utilised the combinations of saturation(S) component based on HSI colour model, thresholding, 5x5 median filter and seeded region growing to segment the nucleus of the WBCs. The proposed system was able to produce an average counting sensitivity of 99.09%. Nayak and Sampathila [11] developed a Raspberry Pi 3-based handheld microscopic device which combined with image processing approaches to perform the leukaemia screening in normal and acute leukaemia images. The results exhibited that the proposed system has produced an average counting accuracy of 99.09%.

Devi *et al.* [12] used Otsu's thresholding as it is a good approach to segment the WBCs in acute leukaemia images. The noise in the segmented images was removed through the median filtering process. Finally, feature extraction has been used to extract shape feature for the classification process. The finding by [12] was supported by Khashman and Al-Zgoul [13] that performed Otsu's thresholding for nucleus segmentation process in acute lymphoblastic leukaemia (ALL), acute myelogenous leukaemia (AML), CLL and CML images. Then, filtering and object elimination was performed to remove the noise after segmentation process. The proposed segmentation method obtained accuracy of 98.33% for nucleus segmentation.

Patel and Mishra [14] suggested *k*-mean clustering for segmentation of ALL images. Feature extraction has been used to extract shape features from the binary image. Finally, image cleaning has been applied to remove the small background pixels inside the segmented image. However, the resultant performance is not much effective (accuracy = 83.3%) as compared to the [13] method. This is supported by Sivakumar and Ramesh [15] study which utilised the *k*-mean clustering for segmentation of WBCs. Then, the morphological operation has been used to invert the resultant image and defining structuring element of the segmented WBCs image.

Besides, a new approach based on a combination of fuzzy, mathematical morphology and thresholding was proposed by Shitong and Min [16]. The performance shows that this method is good to detect the WBCs compared to other conventional methods. However, the main drawback is the proposed method is unable to separate the nucleus and cytoplasm properly. In contrast, the study by Ghosh *et al.* [17] indicated the fuzzy divergence approach is more effective and accurate for the WBCs detection. This method also applied a few mathematical equations such as Gaussian, Gamma and Cauchy. Interestingly, this technique works efficiently to segment the nucleus, however, failed to extract the cytoplasm.

Based on the previous studies, it is found that most of the researchers have conducted segmentation and WBCs counting on acute leukaemia images. There are a low number of studies which process the chronic leukaemia images. Hence, this paper proposes an automated WBCs counting for chronic leukaemia images for leukaemia screening. One of the main difficulties to segment the chronic leukaemia images is that these images consist of overlapping WBCs.

Previous studies have shown that watershed algorithm has become one of the commonly used methods for separation of overlapping blood cells. For instance, Emad *et al*. [7] used the combinations of the watershed algorithm and optimal thresholding to segment the CLL and normal blood images captured under 100X magnification. The proposed segmentation yields 99.92% accuracy for nucleus segmentation and 99.85% accuracy for cell segmentation. Meanwhile, Mishra *et al.* [18] utilised marker-based watershed transformations to decrease the effect of over-segmentation due to the irregular shape of the cells in ALL images. Arslan *et al.* [19] have proposed an algorithm that uses colour and shapes for segmentation of WBCs present in the ALL bone marrow images. Further, this study relies on combining shape and colour information in a marker-controlled watershed algorithm for segmenting the WBCs.

Therefore, the main contribution of this study is to perform the analysis of watershed segmentation using different distance transforms to find the best technique for separating the overlapping cells. As a comparison, three types of distance transform known as Euclidean, city block and chessboard will be applied. The rest of this paper is as follows: Section 2 discusses a methodology of the proposed approach. Section 3 explains the result and discussion. Finally, section 4 describes the conclusion of this research.

II. METHODOLOGY

This study focuses on performing several image processing techniques to obtain the segmented WBCs of chronic leukaemia images. The first technique is the colour conversion of the original RGB image to HSV colour space. Secondly, the image will undergo through segmentation stage based on Otsu's approach. Next, a noise removal technique is performed to remove small unwanted cells and finally, is to separate the overlapping WBCs. The procedures for segmentation of WBCs are illustrated in Figure 1.

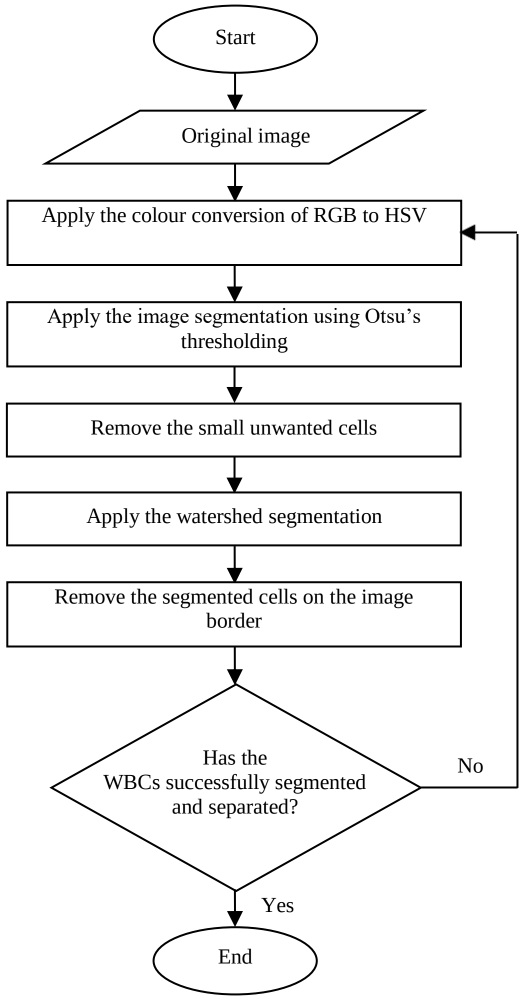


Figure 1: The proposed image processing steps for WBCs segmentation

A. *Image Acquisition*

The first step is to capture the images from the normal and chronic leukaemia blood samples. These blood samples were provided by Hospital Universiti Sains Malaysia (HUSM). In this study, 50 digital images were acquired from CLL, CML and normal blood samples. The slide images were captured at 10X magnification using a computerised Leica DLMA 1200 digital microscope with 800 x 600 resolution and saved in bitmap (*.bmp) format.

B. *Colour Conversion of RGB to HSV Colour Space*

The images captured using a computerised microscope is in RGB image which is complicated for the segmentation process. The RGB image is converted to HSV (hue, saturation, value) colour space for easing the segmentation process. This conversion reduced the colour dimension and image quality. Also, segmentation based on HSV colour space produced a good result compared to RGB colour space [10]. Moreover, for better distinction of cells in the foreground, the saturation component image has been chosen to be segmented as it shows better contrast. For the saturation component, Equation (1) will be interpreted as follows [6]:

$$Saturation = 1 - \frac{3}{R+G+B}\min(R,G,B) \tag{1}$$

C. Image Segmentation using Otsu's Thresholding

After the colour conversion has been performed, the saturation component image is chosen to be segmented by using Otsu's thresholding method. Otsu's thresholding is used to create a binary image from the grey level image. In this study, the saturation component image is segmented into three parts which are for WBCs, red blood cells (RBCs) and background regions. However, after segmentation process using Otsu's thresholding, the result appears some small spots in the background of the segmented image.

Therefore, an arithmetic operation based on area opening is used to remove these small background pixels. This operation allows the eliminating of the unwanted object that is smaller than the structuring element of the desired object. After performing analyses on several normal and chronic leukaemia images, it has been found that the small background pixels may have an area which is less than 80 pixels. Thus, any regions which are less than 80 pixels are considered as non-WBCs and will be eliminated from the image through area opening process.

D. Separation of Cells using Watershed Segmentation with Distance Transform

One of the common features of chronic leukaemia is chronic leukaemia cell is attached to another WBCs. Thus, efficient separation of these overlapping cells could be achieved using watershed segmentation with distance transform. The watershed segmentation technique is based on the concept of topographic representation of image intensity. Furthermore, the watershed segmentation can be divided into several principals of segmentation methods such as discontinuity detection, thresholding and region processing.

Watershed segmentation can be used to segment the connected cells, but need to use distance transform to pre-process the image to make it suitable for watershed segmentation. There are three ways to define the distance transform between black and white pixels which are Euclidean, city block and chessboard and can be computed by using Equations (2), (3) and (4) respectively [20].

$$d_{Euclidean}([i_1, j_1], [i_2, j_2]) = \sqrt{(i_1 - i_2)^2 + (j_1 - j_2)^2} \tag{2}$$

$$d_{Cityblock}([i_1, j_1], [i_2, j_2]) = |i_1 - i_2| + |j_1 - j_2| \tag{3}$$

$$d_{Chessboard}([i_1, j_1], [i_2, j_2]) = \max(|i_1 - i_2|, |j_1 - j_2|) \tag{4}$$

Even though the proposed procedures also utilised watershed segmentation as similar to Emad *et al.* [7], however, the watershed segmentation is applied on chronic leukaemia images that have been captured under 10X magnification (for WBCs counting purpose). As for Emad *et al.* [7], the watershed segmentation has been applied to chronic leukaemia images that have been captured under 100X magnification (for diagnosis purpose). Finally, any segmented WBCs that are attached to the image border is removed to obtain the accurate number of the WBCs.

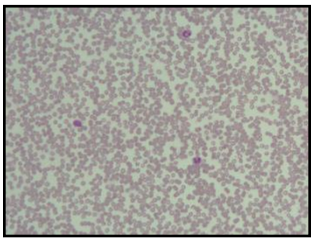
E. Evaluation of White Blood Cells Counting

The WBCs counting process is performed to determine the applicability of watershed segmentation using different distance transform methods for separating the overlapping WBCs. In this study, the WBCs counting is performed by using the region growing algorithm. This algorithm labels the WBCs according to their order in the image and then determines the total number of the WBCs based on the labelled image. Detail descriptions of the procedures to count the number of the WBCs using region growing algorithm can be found in [10].To evaluate the performance of each distance transform method for WBCs counting, the percentage accuracy of the counted WBCs is determined based on the correct counting percentage which is defined as below [21]:

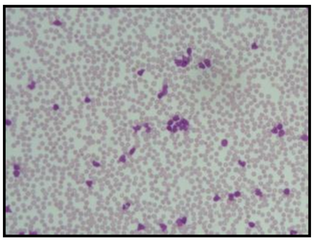
$$Accuracy = \left(1 - \frac{Manual\_Count - System\_Count}{Manual\_Count}\right) * 100\% \tag{5}$$

III. Result and Discussion

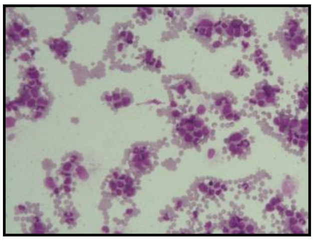
In this study, the proposed image segmentation procedure has been applied and tested on 50 images that have been captured from CLL, CML and normal blood samples. Comparisons of watershed segmentation by using Euclidean, city block and chessboard have been made to measure the performance of each distance transform in separating the overlapping WBCs. The original normal blood, CLL and CML images are presented in Figures 2(a), (b), (c) and (d), respectively.



(a) Normal blood



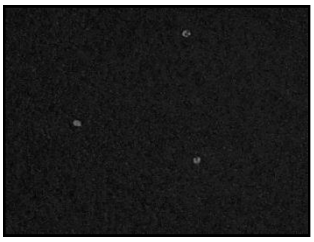
(b) CLL



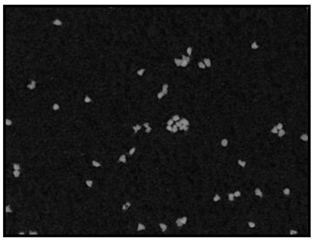
(c) CML

Figure 2: Original images of normal, CLL and CML

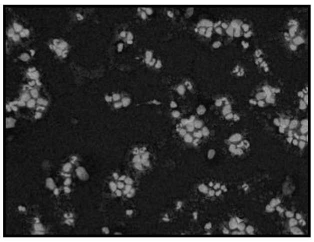
At first, colour conversion of RGB to HSV colour space has been performed to make the segmentation process to be easier. Figure 3 represents the saturation component image that has been extracted from the original blood image. Next, Otsu's thresholding has been applied to segment the WBCs from the background region. Figure 4 shows the resultant images obtained after thresholding the saturation component images. These images include small background pixels which have not been removed through thresholding technique. To obtain a clean segmented image, the small background pixels which are fewer than 80 pixels have been removed through arithmetic operation based area opening technique.



(a) Normal

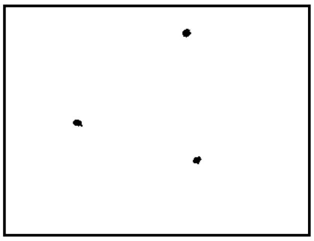


(b) CLL

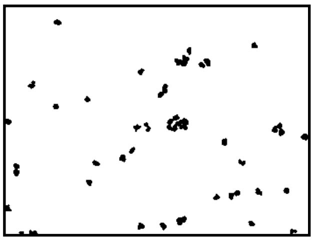


(c) CML

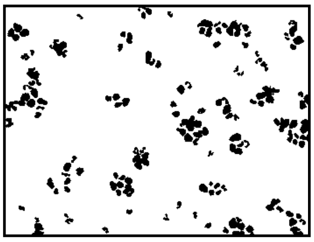
Figure 3: Saturation component images of normal, CLL and CML



(a) Normal



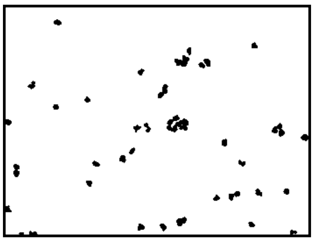
(b) CLL



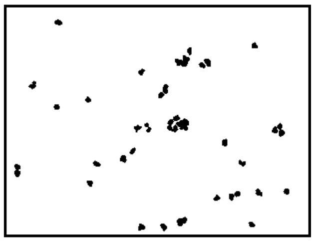
(c) CML

Figure 4: Otsu's images of normal, CLL and CML

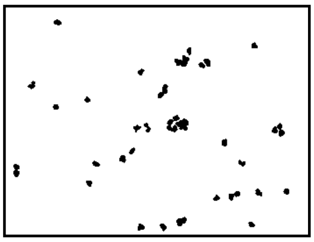
One of the main difficulties to segment the chronic leukaemia images is that these images consist of overlapping WBCs. The overlapping WBCs need to be separated to obtain accurate result during the WBCs counting process. Therefore, watershed segmentation with Euclidean, city block and chessboard distance transforms has been applied to separate the overlapping WBCs. Finally, after separation of WBCs, noise removal at the image border is required to get a better accuracy of the counted WBCs. The final results obtain after performing the watershed segmentation, and clear border process for CLL and CML images are presented in Figures 5 and 6, respectively. From analyses of distance transforms, it can be seen that most of the overlapping WBCs have successfully been separated by using Euclidean as compared to others.



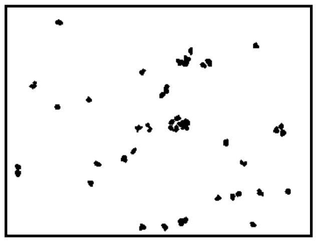
(a) Before separating the cells



(b) Euclidean

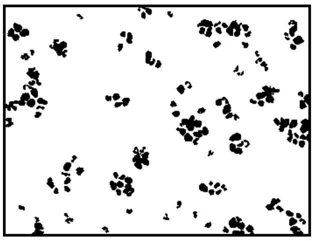


(c) Chessboard

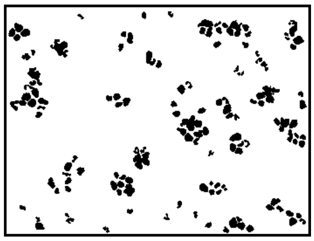


(d) City block

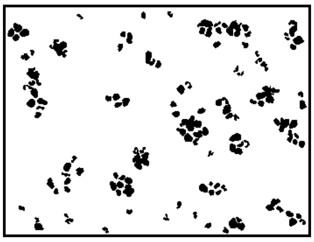
Figure5: Results of watershed segmentation on CLL image using different distance transforms



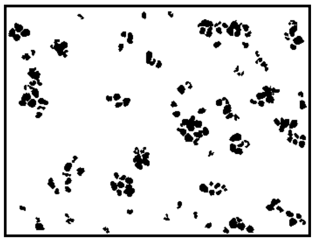
(a) Before separating the cells



(b) Euclidean



(c) Chessboard



(d) City block

Figure 6: Results of watershed segmentation on CML image using different distance transforms

The WBCs counting process has been carried on to validate the performance of watershed segmentation using different distance transforms. The results of WBCs count for normal, CLL and CML images are presented in Tables 1, 2 and 3, respectively. Meanwhile, Table 4 presents the average WBCs count for 50 blood images. From Table 4, the Euclidean distance transform has proven to be the best with average WBCs count of 99.81% as compared to the city block (98.78%) and chessboard (91.09%). The result provided by Euclidean distance transform is slightly higher compared to the results provided by Nayak and Sampathila [11] with average counting accuracy of 99.09%. This result is supported by qualitative analysis as shown in Figures 5 and 6 in which watershed segmentation based on Euclidean can separate most of the overlapping cells, hence increasing the number of WBCs count.

Table 1
WBCs Count for Normal Blood Images

Normal Blood Images	Manual Count	Euclidean	Chessboard	City Block
1	5	6	6	7
2	1	1	1	1
3	2	2	2	2
4	3	3	3	3
5	7	8	7	8
6	4	4	4	4
7	3	3	3	3
8	3	3	3	3
9	2	3	3	3
10	1	1	1	1
11	4	4	4	4
12	3	3	3	3
13	1	3	3	3
14	2	2	2	2
15	0	0	0	0
16	4	4	4	4
Total	45	50	49	51

Table 2
WBCs Count for CLL Images

CLL Images	Manual Count	Euclidean	Chessboard	City Block
1	22	22	22	22
2	26	26	24	25
3	31	31	30	31
4	31	31	31	31
5	32	31	29	31
6	28	27	25	28
7	42	42	40	42
8	44	43	42	43
9	36	36	36	36
10	43	43	40	42
11	31	31	31	31
12	42	40	35	38
13	38	38	35	38
14	46	44	42	44
15	38	39	37	38
16	46	47	43	43
17	48	47	46	46
Total	624	618	588	609

Table 3
WBCs Count for CML images

CML Images	Manual Count	Euclidean	Chessboard	City Block
1	125	123	113	122
2	93	93	82	90
3	99	99	93	101
4	95	95	85	89
5	107	106	96	107
6	138	135	126	131
7	109	110	103	108
8	100	99	90	98
9	116	117	109	116
10	128	129	114	128
11	130	130	113	128
12	147	147	122	144
13	109	109	102	111
14	109	109	97	108
15	140	140	122	138
16	97	98	80	98
17	115	114	108	117
Total	1957	1953	1755	1934

Table 4
Average WBCs Count for 50 Images

Detail	Manual Count	Euclidean	Chessboard	City Block
Total WBCs	2626	2621	2392	2594
Accuracy (%)	--	99.81	91.09	98.78

IV. CONCLUSION

In this study, an automated WBCs counting with analysis of watershed segmentation for the screening of chronic leukaemia images has been presented. Analyses of watershed segmentation using Euclidean, city block and chessboard distance transforms have been made to find the best technique for separating overlapping cells. To segment and perform the WBCs counting process, the blood images have undergone several image processing techniques such as colour conversion, image segmentation using Otsu's method, noise removal and finally separation of overlapping WBCs based on watershed segmentation. To prove the effectiveness, the percentage accuracy of the WBCs count has been calculated by comparing the result of the manual count with automated WBCs count. Based on the average counting accuracy, the watershed segmentation based on Euclidean distance transform gives better performance (99.81%) as compared to the city block (91.09%) and chessboard (98.78%).

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Hospital Universiti Sains Malaysia (HUSM) for providing the leukaemia blood samples. We also would like to acknowledge Universiti Malaysia Perlis for providing the financial support under Short Term Grant Scheme (Grant No. 9001-00514).

REFERENCES

[1] Z. A. Omar, Z. M. Ali, and N. S. I. Tamin, *Malaysian Cancer Statistics: Data and Figure Peninsular Malaysia 2006*. Malaysia: National Cancer Registry, 2006.

[2] G.C.C. Lim, "Overview of cancer in Malaysia," *Japanese Journal of Clinical Oncology*, vol. 32, pp. 37–42, Feb. 2002.

[3] M. Ab Azizah, I.T. Nor Saleha, A. Noor Hashimah, Z.A. Asmah, W. Mastulu, *Malaysian National Cancer Registry Report 2007-2011*, 2016.

[4] G.P.M. Priyankara, O.W. Seneviratne, R.K.O.H. Silva, W.V.D. Soysa, and C.R.D. Silva, *An Extensible Computer Vision Application for Blood Cell Recognition and Analysis*. Sri Lanka: University of Moratuwa, 2006.

[5] C. Reta, L. Altamirano, J.A. Gonzalez, R. Diaz, and J.S. Guichard, "Segmentation of bone marrow cell images for morphological classification of acute leukaemia," in *2010 Proceedings of the Twenty-Third International Florida Artificial Intelligence Research Society Conference*, pp. 86–91.

[6] A.S. Abdul Nasir, M.Y. Mashor, and H. Rosline, "Unsupervised colour segmentation of white blood cell for acute leukaemia images," in *2011 IEEE International Conference on Imaging Systems and Techniques*, pp. 142–145.

[7] E.A. Mohammed, B.H. Far, C. Naugler, and M.M. Mohamed, "Application of support vector machine and k-means clustering algorithms for robust chronic lymphocytic leukaemia color cell segmentation," *2013 IEEE 15th International Conference on e-Health Networking, Applications & Services*, pp. 622–626.

[8] S. Agaian, M. Madhukar, and A.T. Chronopoulos, "Automated screening system for acute myelogenous leukaemia detection in blood microscopic images," *IEEE Systems Journal*, vol. 8, pp. 995–1004, Sep. 2014.

[9] C. Di Ruberto, A. Loddo, and L. Putzu, “A leucocytes count system from blood smear images,” *Machine Vision and Applications*, vol. 27, pp. 1151–1160, Nov. 2016.

[10] N.H. Harun, M.Y. Mashor, H.N. Lim, and R. Hassan, “Automated white blood cells counting system for acute leukaemia based on blood images,”*Jurnal Teknologi*, vol. 78, pp. 91–98, Jan. 2016.

[11] S. K. Nayak and N. Sampathila, “Development of a protocol for screening leukaemia from the microscopic images acquired from blood smear,”*IJRET: International Journal of Research in Engineering and Technology*, vol. 5, pp. 8–11, 2016.

[12] S.D. Devi, R. Sharada, R. Shankari, T. Tamilarasi, and G. Priya, “Automatic Diagnosis of Acute Lymphoblastic Leukemia Using Duplex Method,”*International Journal of Healthcare Sciences*, vol. 5, pp. 14–21, 2017.

[13] A. Khashman and E. Al-zgoul, “Image segmentation of blood cells in leukaemia patients,”*Recent Advances in Computer Engineering and Applications*, pp. 104–109, Jan. 2010.

[14] N. Patel, A. Mishra, “Automated leukaemia detection using microscopic images,”*Procedia Computer Science*, vol. 58, pp. 635–642, 2015.

[15] S. Sivakumar and S. Ramesh, “Automatic white blood cell segmentation using k means clustering,”*International Journal of Science and Engineering Research*, vol. 3, 2015.

[16] W. Shitong and W. Min, “A new detection algorithm (NDA) based on fuzzy cellular neural networks for white blood cell detection,”*IEEE Trans. Inf. Technol. Biomed*, vol. 10, pp. 5–10, 2006.

[17] M. Ghosh, D. Das, C. Chakraborty, and A.K. Ray, “Automated leukocyte recognition using fuzzy divergence,”Micron, vol. 41, pp. 840–846, 2010.

[18] S. Mishra, B. Majhi, P. K. Sa, and L. Sharma, “Gray level co-occurrence matrix and random forest based acute lymphoblastic leukaemia detection,”*Biomedical Signal Processing and Control*, vol. 33, pp. 272–280, Mar. 2017.

[19] S. Arslan, E. Ozyurek, and C. Gunduz-Demir,“A color and shape based algorithm for segmentation of white blood cells in peripheral blood and bone marrow images,”*Cytometry Part A*, vol. 85, pp. 480–490, Jun. 2014.

[20] Q.C.Q. Chen, X.Y.X. Yang, and E.M. Petriu, “Watershed segmentation for binary images with different distance transforms,” in *Proceedings of the 3rd IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications*, pp. 111–116.

[21] N. Theera-Umpon and P.D. Gader, “Training neural networks to count white blood cells via a minimum counting error objective function,” in *2000. Proceedings. 15th International Conference on Pattern Recognition*, pp. 299–302.